Amendments to the Claims

Please amend the claims as follows:

1. (currently amended) A method of identifying a relationship between one or more candidate biomolecules and multiple ~~one-or-more~~ reference biomolecules, the method comprising: (a) inputting to a computer a query set describing the one or more candidate biomolecules; (b) comparing the query set with a target database describing the multiple ~~one-or-more~~ reference biomolecules, wherein the multiple ~~one-or-more~~ reference biomolecules are grouped into one or more buckets, and wherein the ~~one-or-more~~ reference biomolecules of each bucket share a common property; (c) counting a number of matches between each query set and each bucket of the target database; and (d) statistically analyzing each match, wherein the presence of a statistically significant match identifies a relationship between the query set and a bucket of the target database.

2. (original) The method of claim 1, wherein the query set comprises one or more sequences.

3. (original) The method of claim 2, wherein the one or more sequences are selected from the group consisting of a DNA sequence, an RNA sequence, and a protein sequence.

4. (original) The method of claim 2, wherein the one or more sequences are extracted from one genetic region.

5. (currently amended) The method of claim 1, wherein the one or more candidate biomolecules and the multiple ~~one-or-more~~ reference biomolecules are all selected from the group consisting of proteins, nucleic acids, and small molecules.

6. (original) The method of claim 1, wherein the comparing comprises employing an equivalence algorithm based on identity of name, accession, or other identifier associated with biomolecule.

7. (original) The method of claim 1, wherein the comparing comprises employing a BLAST-based algorithm to identify similarities or identities in two or more sequences.

8. (original) The method of claim 1, wherein the counting comprises applying one or more principles chosen from the group consisting of: (a) each query set candidate sequence can match at most one reference sequence in any given bucket; (b) each query set candidate sequence can possess a match in one or more different buckets; and (c) once a candidate sequence in the query set matches a specific bucket reference sequence in the target database, any subsequent matches of that same candidate sequence to other reference sequences in that bucket do not increase the match count for the bucket.

9. (original) The method of claim 1, wherein the statistically analyzing comprises computing one or more statistics for each bucket.

10 .(original) The method of claim 8, further comprising sorting the one or more statistics by increasing or decreasing significance.

11. (original) The method of claim 1, further comprising outputting a webpage with results of the statistical analysis, the webpage comprising one or more hyperlinks.

12. (currently amended) A computer-readable storage device embodying a program of instructions executable by a computer to perform method steps for identifying a relationship between one or more candidate biomolecules and multiple ~~one or more~~ reference biomolecules, the method steps comprising: (a) inputting to a computer a query set describing one or more candidate biomolecules; (b) comparing the query set with a target database describing multiple ~~one or more~~ reference biomolecules, the multiple ~~one or more~~ reference biomolecules of the target database grouped into one or more buckets, wherein the ~~one or more~~ reference biomolecules of each bucket share a common property; (c) counting a number of matches between each query set and each bucket of the target database; and (d) statistically analyzing each match, wherein the presence of a

3

statistically significant match identifies a relationship between a query set and one or more buckets of a target database.

13. (original) The computer-readable storage device of claim 12, wherein the query set comprises one or more candidate sequences.

14. (original) The computer-readable storage device of claim 13, wherein the one or more candidate sequences are selected from the group consisting of a DNA sequence, an RNA sequence, and a protein sequence.

15. (original) The computer-readable storage device of claim 13, wherein the one or more candidate sequences are extracted from one genetic region.

16. (currently amended) The computer-readable storage device of claim 12, wherein the one or more candidate biomolecules and the multiple one or more reference biomolecules are all selected from the group consisting of proteins, nucleic acids, and small molecules.

17. (original) The computer-readable storage device of claim 12, wherein the comparing comprises employing a BLAST-based algorithm to identify similarities or identities in two or more sequences.

18. (original) The computer-readable storage device of claim 12, wherein the comparing comprises employing a equivalence algorithm based on identity of name, accession, or other identifier associated with biomolecule.

19. (original) The computer-readable storage device of claim 12, wherein the counting comprises applying one or more principles chosen from the group consisting of: (a) each query set candidate sequence can match at most one reference sequence in any given bucket; (b) each query set candidate sequence can possess a match in one or more different buckets; and (c) once a candidate sequence in the query set matches a specific

bucket reference sequence in the target database, any subsequent matches of that same candidate sequence to other reference sequences in that bucket do not increase the match count for the bucket.

20. (original) The computer-readable storage device of claim 12, wherein the statistically analyzing comprises computing one or more statistics for each match.

21. (original) The computer-readable storage device of claim 20, further comprising sorting the one or more statistics by increasing or decreasing significance.

22. (original) The computer-readable storage device of claim 12, further comprising outputting a webpage with results of the statistically analyzing, the webpage comprising one or more hyperlinks.

23. (currently amended) A method of identifying a relationship between two or more region sets, each region set describing one or more candidate biomolecules, and a target database describing multiple ~~one or more~~ reference biomolecules grouped into one or more buckets, the method comprising: (a) providing a query set describing two or more region sets, each region set comprising one or more candidate biomolecule sequences extracted from one region; (b) comparing the query set with target database sequences describing multiple ~~one or more~~ reference biomolecule sequences, wherein the target database sequences are grouped into one or more buckets, and wherein the ~~one or more~~ reference biomolecules of each bucket share a common property; (c) counting a number of matches between each query set and each bucket of the target database; and (d) statistically analyzing each match, wherein the presence of a statistically significant match identifies a relationship between the query set and the bucket of the target database.

24. (original) The method of claim 23, wherein the one or more biomolecule sequences are selected from the group consisting of protein sequences and nucleic acid sequences.

25. (original) The method of claim 24, wherein the nucleic acid sequences are selected from the group consisting of a DNA sequence and an RNA sequence.

26. (original) The method of claim 23, wherein the comparing comprises employing a equivalence algorithm based on identity of name, accession, or other identifier associated with biomolecule. 2

27. (original) The method of claim 23, wherein the comparing comprises employing a BLAST-based algorithm to identify similarities or identities in two or more sequences.

28. (original) The method of claim 23, wherein the counting comprises applying one or more principles chosen from the group consisting of: (a) each query set candidate sequence can match at most one reference sequence in any given bucket; (b) each query set candidate sequence can possess a match in one or more different buckets; and (c) once a candidate sequence in the query set matches a specific bucket reference sequence in the target database, any subsequent matches of that same candidate sequence to other reference sequences in that bucket do not increase the match count for the bucket.

29. (original) The method of claim 23, wherein the statistically analyzing comprises computing one or more statistics for each match.

30. (original) The method of claim 29, further comprising sorting the one or more statistics by increasing or decreasing significance.

31. (original) The method of claim 30, further comprising further comprising outputting a webpage with results of the statistically analyzing, the webpage comprising one or more hyperlinks.

32. (original) The method of claim 23, the method further comprising: (a) constructing a plurality of replicates of the one or more query sets; (b) modeling the

replicates at random chromosomal locations to form a random location data set; (c) processing the random location data set by following the steps of claim 23; (d) quantifying the number of times each match is found to surpass a predetermined threshold to form a statistically significant set of random location matches; and (e) comparing the statistically significant set of random location matches to the statistically significant relationship of claim 23.

33. (currently amended) A computer-readable storage device embodying a program of instructions executable by a computer to perform method steps for identifying a relationship between two or more region sets, each region set ~~each region set~~ describing one or more candidate biomolecules, and a target database describing multiple ~~one or more~~ reference biomolecules grouped into one or more buckets, the method steps comprising: (a) providing a query set describing two or more region sets, each region set comprising one or more candidate biomolecule sequences extracted from one genetic region; (b) comparing the query set with target database sequences describing multiple ~~one or more~~ reference biomolecule sequences, wherein the target database sequences grouped into one or more buckets, and wherein the ~~one or more~~ reference biomolecules of each bucket share a common property; (c) counting a number of matches between each query set and each bucket of the target database; and (d) statistically analyzing each match, wherein the presence of a statistically significant match identifies a relationship between the query set and the bucket of the target database.

34. (currently amended) The computer-readable storage device of claim 33, wherein the one or more candidate biomolecule sequences and the multiple ~~one or more~~ reference biomolecules sequences are all selected from the group consisting of protein sequences and nucleic acid sequences.

35. (original) The computer-readable storage device of claim 34, wherein the nucleic acid sequences are selected from the group consisting of a DNA sequence and an RNA sequence.

36. (original) The computer-readable storage device of claim 33, wherein the comparing comprises employing a BLAST-based algorithm to identify similarities or identities in two or more sequences.

37. (original) The computer-readable storage device of claim 33, wherein the counting comprises applying one or more principles chosen from the group consisting of: (a) each query set candidate sequence can match at most one reference sequence in any given bucket; (b) each query set candidate sequence can possess a match in one or more different buckets; and (c) once a candidate sequence in the query set matches a specific bucket reference sequence in the target database, any subsequent matches of that same candidate sequence to other reference sequences in that bucket do not increase the match count for the bucket.

38. (original) The computer-readable storage device of claim 33, wherein the statistically analyzing comprises computing one or more statistics for each match.

39. (original) The computer-readable storage device of claim 38, further comprising sorting the one or more statistics by increasing or decreasing significance.

40. (original) The computer-readable storage device of claim 39, further comprising outputting a webpage with results of the statistically analyzing, the webpage comprising one or more hyperlinks.

41. (original) The computer-readable storage device of claim 33, the method steps further comprising: (a) constructing a plurality of replicates of the one or more query sets; (b) modeling the replicates at random chromosomal locations to form a random location data set; (c) processing the random location data set by following the steps of claim 33; (d) quantifying the number of times each match is found to surpass a predetermined threshold to form a statistically significant set of random location matches; and (e) comparing the statistically significant set of random location matches to the statistically significant relationship of claim 33.

42. -66. Cancelled.